3. Adler, S. *et al. Arch. Toxicol.* **85**, 367–485 (2011).
4. Krewski, D. *et al. J. Toxicol. Environ. Health B Crit. Rev.* **13**, 51–138 (2010).
5. Organization for economic cooperation and development. Report of the workshop on using mechanistic information in forming chemical categories. (no. 138, series on testing and assessment) (OECD, Crystal City, 2011).
6. Kavlock, R., *et al. Chem. Res. Toxicol.* **25**, 1287–1302 (2012).
7. Hall, B. *et al. Food Chem. Toxicol.* **45**, 2097–2108 (2007).
8. Adeleye, Y. *et al. Toxicology* available online, doi:10.1016/j.tox.2014.02.007 (25 February 2014).

RNA sequencing *in situ*

Paul Ginart & Arjun Raj

Using a cell as an RNA sequencing chip enables spatial analyses of the transcriptome at subcellular resolution.

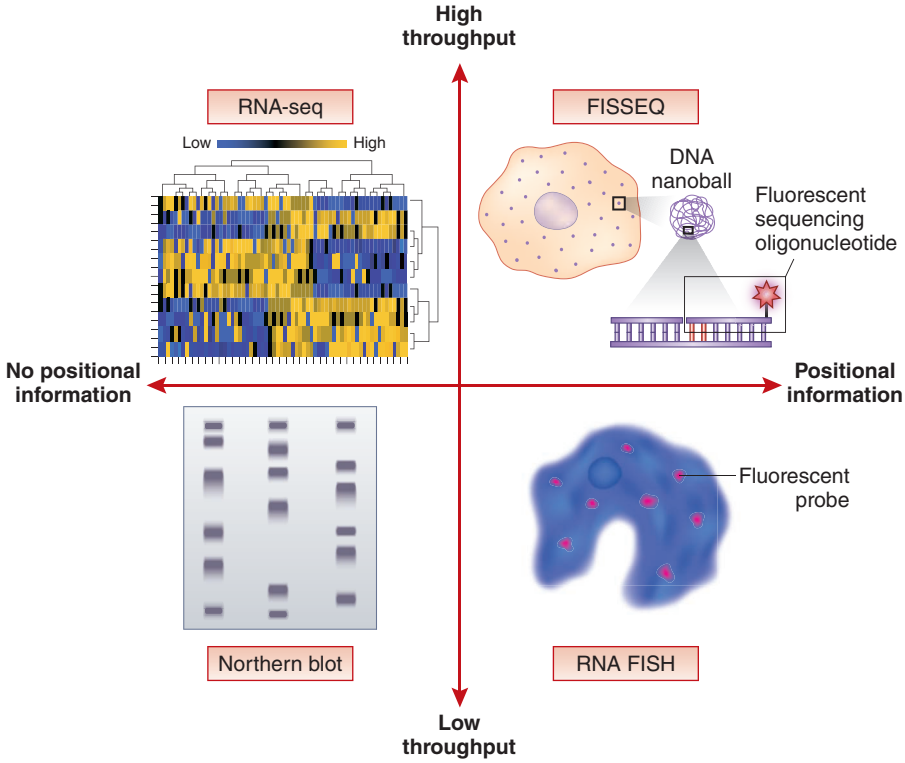
A complete accounting of gene expression in an organism would provide the exact sequence and location of every transcript in every cell. This goal is a long way off, but significant progress has been made through decades of technology development in sequencing and imaging. On one hand, RNA sequencing (RNA-seq) has made it possible to measure gene expression on a genome-wide scale, typically in cell populations[1]. On the other hand, RNA fluorescence *in situ* hybridization (FISH) has advanced to the point where we can detect individual RNA molecules in single cells through a microscope, gaining accurate, absolute single-cell expression data together with information on RNA location down to the subcellular level, but usually for just a few genes at a time[2,3]. Writing in *Science*, Lee *et al.*[4] report an exciting first step toward combining some of the best aspects of both methods to sequence RNA in single cells *in situ*. Their approach, called fluorescence *in situ* RNA sequencing (FISSEQ), effectively treats the cell as a sequencing chip while preserving spatial information about RNA location.

The authors begin by fixing cells, reverse transcribing the RNA *in situ* and copying the resulting cDNAs by rolling-circle amplification to generate single-stranded DNA 'nanoballs' (**Fig. 1**). Each nanoball, about 200–400 nm in diameter, consists of many cDNA copies of an original RNA sequence. The large number of repeats means that the nanoballs can withstand repeated sequencing reactions, allowing the authors to analyze them inside cells using sequencing by oligonucleotide ligation and detection (SOLiD). In this sequencing method, fluorescent probes are sequentially hybridized to the nanoballs, with a color corresponding to a pair of adjacent bases, and the sequence of colors is used to determine the sequence of the transcript. Because this process involves repeated imaging of the sample, careful image analysis is needed to reconstruct the sequences.

Working with primary human fibroblasts, Lee *et al.*[4] detect RNA from ~12,500 DNA nanoballs representing transcripts from ~4,000 genes. Reads are 27 bp long on average, with a median per-base error rate of 0.64%. The authors also show that the approach can be used for differential expression analysis in a wound-healing assay. Importantly, FISSEQ captures many fibroblast-specific gene expression markers, demonstrating its potential utility in unbiased cell-type identification.

This work represents a technical *tour de force* that required the solution of several chemical and computational challenges. Performing enzyme-catalyzed nucleic acid manipulations *in situ* has proven difficult, and although other groups have demonstrated some steps in the FISSEQ method, such as *in situ* rolling-circle amplification, no one, to our knowledge, has attempted anything approaching the complexity of sequencing inside the cell itself.

From the computational perspective, FISSEQ depends on sophisticated image analysis, including robust identification and validation of signals within the image. Indeed, a key realization in the paper is that the sequencing provides an internal consistency check. The signal from each fluorescent spot in an image corresponds to a sequencing read, which can be mapped against a reference transcriptome. Spots that generate mappable reads are more likely to represent an



Debbie Maizels/Nature Publishing Group

Figure 1 Choosing the proper assay for gene expression involves trade-offs on parameters such as throughput or positional resolution. Northern blots assay a few genes at a time and average over bulk samples. RNA-seq provides tremendous throughput, but positional information is lost. RNA FISH gives subcellular positional resolution of gene expression but currently cannot be highly multiplexed. FISSEQ, although it still has technical hurdles to overcome, stands as the most comprehensive gene expression assay, combining the breadth of RNA-seq with the subcellular resolution of RNA FISH.

Paul Ginart and Arjun Raj are at the University of Pennsylvania, Philadelphia, Pennsylvania, USA.
e-mail: arjunrajlab@gmail.com

RNA in the cell, whereas nonmapping reads are indicative of background noise, autofluorescence or debris.

Another issue is the enormous range in intracellular RNA abundance, which pits the need for a large number of reads against the fundamental optical limitations on the number of spot-like signals that can be reliably detected within a cell. To solve this problem, the authors develop 'partition sequencing', a clever method in which they extend the sequencing primers by a fixed number of bases so that only some of the RNA molecules present in the cell are sequenced. This allows them to tune the number of sequence readouts to avoid saturating the imaging field with too many targets, facilitating robust yet broad sampling of the RNA in a cell.

In a step toward studying multicellular systems, the authors use FISSEQ to detect RNA transcription in a whole-mount *Drosophila* embryo as well as in a mouse embryo and in brain sections. They have not yet sequenced these samples but have shown that, in principle, FISSEQ can be extended to study more complicated biological systems.

FISSEQ has the potential to transform several areas of research. In development, for example, the spatial organization of transcription in a multicellular organism is of critical importance. Many disease processes may also involve spatially heterogeneous gene expression, such as that seen in different areas of tumors. FISSEQ analysis of a tumor biopsy may be useful for guiding treatment programs, in which multiple drugs might be chosen for different components of the tumor.

In its current form, FISSEQ does not provide a complete accounting of the RNA within single cells. Rather, it produces a spatially localized subsampling of the transcriptome, and comparisons with RNA-seq and microarray data suggest that FISSEQ may miss lower-abundance transcripts. Indeed, Lee *et al.*[4] detect a large number of transcripts by only a single read per cell. Whether this level of detection is sufficient depends on the scientific question being asked. As the authors note, it may be enough to distinguish different cell types owing to the sheer number of genes detected, in contrast to lower-throughput methods such as RNA FISH. However, accurately quantifying cell-to-cell variability in the expression of a particular gene would be difficult. In light of the rapid recent advances in sequencing technologies and in the imaging and fluid-handling capabilities of current commercial instruments, it will not be surprising if this limitation is addressed in the near future.

At the same time, FISSEQ should have some competition from new variants of RNA-seq and RNA FISH. Several groups have sequenced RNA from single cells, now up to thousands per study[5], and many groups have been steadily improving single-molecule RNA FISH by increasing multiplex capability[6,7], measuring transcriptional structure in the nucleus[8], achieving single-base resolution[9,10] and shortening assay time[11]. Nevertheless, FISSEQ represents an enormous step forward in the ability to measure gene expression in its spatial context, and we look forward to seeing how it will be applied to yield new biological discoveries.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Wang, Z., Gerstein, M. & Snyder, M. *Nat. Rev. Genet.* **10**, 57–63 (2009).
2. Femino, A.M., Fay, F.S., Fogarty, K. & Singer, R.H. *Science* **280**, 585–590 (1998).
3. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. *Nat. Methods* **5**, 877–879 (2008).
4. Lee, J.H. *et al. Science* **343**, 1360–1363 (2014).
5. Jaitin, D.A. *et al. Science* **343**, 776–779 (2014).
6. Lubeck, E. & Cai, L. *Nat. Methods* **9**, 743–748 (2012).
7. Lubeck, E., Coskun, A.F., Zhiyentayev, T., Ahmad, M. & Cai, L. *Nat. Methods* **11**, 360–361 (2014).
8. Levesque, M.J. & Raj, A. *Nat. Methods* **10**, 246–248 (2013).
9. Levesque, M.J., Ginart, P., Wei, Y. & Raj, A. *Nat. Methods* **10**, 865–867 (2013).
10. Hansen, C.H. & van Oudenaarden, A. *Nat. Methods* **10**, 869–871 (2013).
11. Shaffer, S.M., Wu, M.-T., Levesque, M.J. & Raj, A. *PLoS ONE* **8**, e75120 (2013).

Research Highlights

Papers from the literature selected by the Nature Biotechnology editors.
(Follow us on Twitter, @NatureBiotech #nbtHighlight)

Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development
Bendall, S.C. *et al. Cell* **157**, 714–725 (2014)

Universal computing by DNA origami robots in a living animal
Amir, Y. *et al. Nat. Nanotechnol.* **9**, 353–357 (2014)

Single-cell genomics for dissection of complex malaria infections
Nair, S. *et al. Genome Res.* doi:10.1101/gr.168286.113 (8 May 2014)

Multifunctional wearable devices for diagnosis and therapy of movement disorders
Son, D. *et al. Nat. Nanotechnol.* **9**, 397–404 (2014)

Expression and glycoengineering of functionally active heteromultimeric IgM in plants
Loos, A. *et al. PNAS* **111**, 6263–6268 (2014)